



OIPE

ENTERED

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/889,722A

DATE: 03/27/2002  
TIME: 13:32:53

Input Set : A:\sequence listing.txt  
Output Set: N:\CRF3\03272002\I889722A.raw

4 <110> APPLICANT: Japan Science and Technology Corporation  
6 <120> TITLE OF INVENTION: Human nucleoprotein having a WW domain and  
7 a polynucleotide encoding the protein  
9 <130> FILE REFERENCE: 09/889,722  
C--> 11 <140> CURRENT APPLICATION NUMBER: US/09/889,722A  
C--> 12 <141> CURRENT FILING DATE: 2001-07-20  
14 <150> PRIOR APPLICATION NUMBER: JP11-332572  
15 <151> PRIOR FILING DATE: 1999-11-24  
17 <160> NUMBER OF SEQ ID NOS: 7  
19 <170> SOFTWARE: PatentIn Ver. 2.0  
21 <210> SEQ ID NO: 1  
22 <211> LENGTH: 704  
23 <212> TYPE: PRT  
24 <213> ORGANISM: Homo sapiens  
26 <400> SEQUENCE: 1  
27 Met Ala Asn Glu Asn His Gly Ser Pro Arg Glu Glu Ala Ser Leu Leu  
28 1 5 10 15  
29 Ser His Ser Pro Gly Thr Ser Asn Gln Ser Gln Pro Cys Ser Pro Lys  
30 20 25 30  
31 Pro Ile Arg Leu Val Gln Asp Leu Pro Glu Glu Leu Val His Ala Gly  
32 35 40 45  
33 Trp Glu Lys Cys Trp Ser Arg Arg Glu Asn Arg Pro Tyr Tyr Phe Asn  
34 50 55 60  
35 Arg Phe Thr Asn Gln Ser Leu Trp Glu Met Pro Val Leu Gly Gln His  
36 65 70 75 80  
37 Asp Val Ile Ser Asp Pro Leu Gly Leu Asn Ala Thr Pro Leu Pro Gln  
38 85 90 95  
39 Asp Ser Ser Leu Val Glu Thr Pro Pro Ala Glu Asn Lys Pro Arg Lys  
40 100 105 110  
41 Arg Gln Leu Ser Glu Glu Gln Pro Ser Gly Asn Gly Val Lys Lys Pro  
42 115 120 125  
43 Lys Ile Glu Ile Pro Val Thr Pro Thr Gly Gln Ser Val Pro Ser Ser  
44 130 135 140  
45 Pro Ser Ile Pro Gly Thr Pro Thr Leu Lys Met Trp Gly Thr Ser Pro  
46 145 150 155 160  
47 Glu Asp Lys Gln Gln Ala Ala Leu Leu Arg Pro Thr Glu Val Tyr Trp  
48 165 170 175  
49 Asp Leu Asp Ile Gln Thr Asn Ala Val Ile Lys His Arg Gly Pro Ser  
50 180 185 190  
51 Glu Val Leu Pro Pro His Pro Glu Val Glu Leu Leu Arg Ser Gln Leu  
52 195 200 205  
53 Ile Leu Lys Leu Arg Gln His Tyr Arg Glu Leu Cys Gln Gln Arg Glu  
54 210 215 220

RECEIVED  
APR 11 2002  
TECH CENTER 1600/2000

## RAW SEQUENCE LISTING

DATE: 03/27/2002

PATENT APPLICATION: US/09/889,722A

TIME: 13:32:53

Input Set : A:\sequence listing.txt

Output Set: N:\CRF3\03272002\I889722A.raw

```

55 Gly Ile Glu Pro Pro Arg Glu Ser Phe Asn Arg Trp Met Leu Glu Arg
56 225                230                235                240
57 Lys Val Val Asp Lys Gly Ser Asp Pro Leu Leu Pro Ser Asn Cys Glu
58                245                250                255
59 Pro Val Val Ser Pro Ser Met Phe Arg Glu Ile Met Asn Asp Ile Pro
60                260                265                270
61 Ile Arg Leu Ser Arg Ile Lys Phe Arg Glu Glu Ala Lys Arg Leu Leu
62                275                280                285
63 Phe Lys Tyr Ala Glu Ala Ala Arg Arg Leu Ile Glu Ser Arg Ser Ala
64                290                295                300
65 Ser Pro Asp Ser Arg Lys Val Val Lys Trp Asn Val Glu Asp Thr Phe
66 305                310                315                320
67 Ser Trp Leu Arg Lys Asp His Ser Ala Ser Lys Glu Asp Tyr Met Asp
68                325                330                335
69 Arg Leu Glu His Leu Arg Arg Gln Cys Gly Pro His Val Ser Ala Ala
70                340                345                350
71 Ala Lys Asp Ser Val Glu Gly Ile Cys Ser Lys Ile Tyr His Ile Ser
72                355                360                365
73 Leu Glu Tyr Val Lys Arg Ile Arg Glu Lys His Leu Ala Ile Leu Lys
74                370                375                380
75 Glu Asn Asn Ile Ser Glu Glu Val Glu Ala Pro Glu Val Glu Pro Arg
76 385                390                395                400
77 Leu Val Tyr Cys Tyr Pro Val Arg Leu Ala Val Ser Ala Pro Pro Met
78                405                410                415
79 Pro Ser Val Glu Met His Met Glu Asn Asn Val Val Cys Ile Arg Tyr
80                420                425                430
81 Lys Gly Glu Met Val Lys Val Ser Arg Asn Tyr Phe Ser Lys Leu Trp
82                435                440                445
83 Leu Leu Tyr Arg Tyr Ser Cys Ile Asp Asp Ser Ala Phe Glu Arg Phe
84                450                455                460
85 Leu Pro Arg Val Trp Cys Leu Leu Arg Arg Tyr Gln Met Met Phe Gly
86 465                470                475                480
87 Val Gly Leu Tyr Glu Gly Thr Gly Leu Gln Gly Ser Leu Pro Val His
88                485                490                495
89 Val Phe Glu Ala Leu His Arg Leu Phe Gly Val Ser Phe Glu Cys Phe
90                500                505                510
91 Ala Ser Pro Leu Asn Cys Tyr Phe Arg Gln Tyr Cys Ser Ala Phe Pro
92                515                520                525
93 Asp Thr Asp Gly Tyr Phe Gly Ser Arg Gly Pro Cys Leu Asp Phe Ala
94                530                535                540
95 Pro Leu Ser Gly Ser Phe Glu Ala Asn Pro Pro Phe Cys Glu Glu Leu
96 545                550                555                560
97 Met Asp Ala Met Val Ser His Phe Glu Arg Leu Leu Glu Ser Ser Pro
98                565                570                575
99 Glu Pro Leu Ser Phe Ile Val Phe Ile Pro Glu Trp Arg Glu Pro Pro
100                580                585                590
101 Thr Pro Ala Leu Thr Arg Met Glu Gln Ser Arg Phe Lys Arg His Gln
102                595                600                605
103 Leu Ile Leu Pro Ala Phe Glu His Glu Tyr Arg Ser Gly Ser Gln His

```

## RAW SEQUENCE LISTING

DATE: 03/27/2002

PATENT APPLICATION: US/09/889,722A

TIME: 13:32:53

Input Set : A:\sequence listing.txt

Output Set: N:\CRF3\03272002\I889722A.raw

```

104      610      615      620
105 Ile Cys Lys Lys Glu Glu Met His Tyr Lys Ala Val His Asn Thr Ala
106 625      630      635      640
107 Val Leu Phe Leu Gln Asn Asp Pro Gly Phe Ala Lys Trp Ala Pro Thr
108      645      650      655
109 Pro Glu Arg Leu Gln Glu Leu Ser Ala Ala Tyr Arg Gln Ser Gly Arg
110      660      665      670
111 Ser His Ser Ser Gly Ser Ser Ser Ser Ser Ser Ser Glu Ala Lys Asp
112      675      680      685
113 Arg Asp Ser Gly Arg Glu Gln Gly Pro Ser Arg Glu Pro His Pro Thr
114      690      695      700
117 <210> SEQ ID NO: 2
118 <211> LENGTH: 2112
119 <212> TYPE: DNA
120 <213> ORGANISM: Homo sapiens
122 <400> SEQUENCE: 2
123 atggccaatg agaatcacgg cagcccccg gaggaagcgt cctgtctgag tcaactcccca 60
124 ggtacctcca atcagagcca gccctgttct ccaaagccaa tccgcctggt tcaggacctc 120
125 ccagaggagc tgggtgcatgc aggctgggag aagtgtctga gccggaggga gaatcgtccc 180
126 tactacttca accgattcac caaccagtcc ctgtgggaga tgcccgtgct ggggcagcac 240
127 gatgtgattt cggacccttt ggggctgaat cgcaccccac tgcccgaaga ctcaagcttg 300
128 gtggaaactc ccccggtcga gaacaagccc agaaagcggc agctctcgga agagcagcca 360
129 agcggcaatg gtgtgaagaa gcccagatt gaaatcccag tgacaccac aggccagtcg 420
130 gtgcccagct ccccagtat cccaggaacc ccaacgctga agatgtggg tacgtcccct 480
131 gaagataaac agcaggcagc tctcctacga ccaactgagg tctactggga cctggacatc 540
132 cagaccaatg ctgtcatcaa gcaccggggg ccttcagagg tgctgcccc gcaccccgaa 600
133 gtggaactgc tccgtcttca gctcatcctg aagcttcggc agcactatcg ggagctgtgc 660
134 cagcagcgag agggcattga gctccacgg gagtctttca accgctggat gctggagcgc 720
135 aaggtggtag acaaaggatc tgaccccctg ttgcccagca actgtgaacc agtcgtgtca 780
136 ccttccatgt ttcgtgaaat catgaacgac attcctatca ggttatcccg aatcaagttc 840
137 cgggaggaag ccaagcgcct gctctttaaa tatgoggagg ccgccaggcg gctcatcgag 900
138 tccaggagtg catcccctga cagtaggaag gtggtcaaat ggaatgtgga agacaccttt 960
139 agctggcttc ggaaggacca ctccagcctc aaggaggact acatggatcg cctggagcat 1020
140 ctgoggaggc agtgtggccc ccacgtctcg gccgcagcca aggactccgt ggaaggcatc 1080
141 tgcagtaaga tctaccacat ctccctggag tacgtcaaac ggatccgaga gaagcacctt 1140
142 gccatcctca aggaaaacaa catctcagag gaggtggagg cccctgaggt ggagccccgc 1200
143 ctagtgtact gctacccagt ccggctggct gtgtctgcac cgcccatgcc cagcgtggag 1260
144 atgcacatgg agaacaacgt ggtctgcac cggtataagg gagagatggt caaggtcagc 1320
145 cgcaactact tcagcaagct gtggctcctt taccgctaca gctgcattga tgactctgcc 1380
146 tttgagaggt tctgcccccg ggtctgggtg cttctccgac ggtaccagat gatgttcggc 1440
147 gtgggcctct acgaggggac tggcctgcag ggatcgctgc ctgtgcatgt ctttgaggcc 1500
148 ctccaccgac tctttggcgt cagcttcgag tgcttcgcct caccctcaa ctgctacttc 1560
149 cgccagtact gttctgcctt ccccgacaca gacggctact ttggctcccg cgggccttgc 1620
150 ctagactttg ctccactgag tggttcatth gaggccaaac ctcccttctg cgaggagctc 1680
151 atggatgcca tgggtcttca ctttgagaga ctgcttgaga gctcaccgga gccctgtcc 1740
152 ttcacgtgtg tcatccctga gtggcgggaa cccccaacac cagcgtcac ccgcatggag 1800
153 cagagccgct tcaaacgcca ccagttgatc ctgcctgcct ttgagcatga gtaccgcagt 1860
154 ggctcccagc acatctgcaa gaaggaggaa atgcactaca aggccgtcca caacacggct 1920
155 gtgctcttcc tacagaacga ccctggcttt gccaaagtgg cgccgacgcc tgaacggctg 1980

```

## RAW SEQUENCE LISTING

DATE: 03/27/2002

PATENT APPLICATION: US/09/889,722A

TIME: 13:32:53

Input Set : A:\sequence listing.txt

Output Set: N:\CRF3\03272002\I889722A.raw

```

156 caggagctga gtgctgccta ccggcagtcg gccgcagcc acagctctgg ttcttctca 2040
157 tcgtctctct cggaggccaa ggaccgggac tcgggccgtg agcaggggtcc tagccgcgag 2100
158 cctcacccca ct 2112
161 <210> SEQ ID NO: 3
162 <211> LENGTH: 2669
163 <212> TYPE: DNA
164 <213> ORGANISM: Homo sapiens
166 <220> FEATURE:
167 <221> NAME/KEY: CDS
168 <222> LOCATION: (292)..(2406)
170 <400> SEQUENCE: 3
171 acacaagatg gcggcagcgg cgctggggag ggcgagggcg aggcgggcaaa acgggcggtc 60
172 gagcagaacg tgtagccgcg tccccctccag tccgctccgg gcagctgctg atgcaaggaa 120
173 tccccctggg tcccgctccac tccactgctg accagcccat tcgcctgtgc tgagtcttcc 180
174 tgcaggccct tcttgcctc tgtgggaccc tgtgggggtc catccggctg gagaagaaaa 240
175 gcctctcatg ctaacgttgc agaccccgaga gggctcctgt tgggtgtgga g atg gcc 297
176 Met Ala
177 1
178 aat gag aat cac ggc agc ccc cgg gag gaa gcg tcc ctg ctg agt cac 345
179 Asn Glu Asn His Gly Ser Pro Arg Glu Glu Ala Ser Leu Leu Ser His
180 5 10 15
181 tcc cca ggt acc tcc aat cag agc cag ccc tgt tct cca aag cca atc 393
182 Ser Pro Gly Thr Ser Asn Gln Ser Gln Pro Cys Ser Pro Lys Pro Ile
183 20 25 30
184 cgc ctg gtt cag gac ctc cca gag gag ctg gtg cat gca ggc tgg gag 441
185 Arg Leu Val Gln Asp Leu Pro Glu Glu Leu Val His Ala Gly Trp Glu
186 35 40 45 50
187 aag tgc tgg agc cgg agg gag aat cgt ccc tac tac ttc aac cga ttc 489
188 Lys Cys Trp Ser Arg Arg Glu Asn Arg Pro Tyr Tyr Phe Asn Arg Phe
189 55 60 65
190 acc aac cag tcc ctg tgg gag atg ccc gtg ctg ggg cag cac gat gtg 537
191 Thr Asn Gln Ser Leu Trp Glu Met Pro Val Leu Gly Gln His Asp Val
192 70 75 80
193 att tcg gac cct ttg ggg ctg aat gcg acc cca ctg ccc caa gac tca 585
194 Ile Ser Asp Pro Leu Gly Leu Asn Ala Thr Pro Leu Pro Gln Asp Ser
195 85 90 95
196 agc ttg gtg gaa act ccc ccg gct gag aac aag ccc aga aag cgg cag 633
197 Ser Leu Val Glu Thr Pro Pro Ala Glu Asn Lys Pro Arg Lys Arg Gln
198 100 105 110
199 ctc tcg gaa gag cag cca agc ggc aat ggt gtg aag aag ccc aag att 681
200 Leu Ser Glu Glu Gln Pro Ser Gly Asn Gly Val Lys Lys Pro Lys Ile
201 115 120 125 130
202 gaa atc cca gtg aca ccc aca ggc cag tcg gtg ccc agc tcc ccc agt 729
203 Glu Ile Pro Val Thr Pro Thr Gly Gln Ser Val Pro Ser Ser Pro Ser
204 135 140 145
205 atc cca gga acc cca acg ctg aag atg tgg ggt acg tcc cct gaa gat 777
206 Ile Pro Gly Thr Pro Thr Leu Lys Met Trp Gly Thr Ser Pro Glu Asp
207 150 155 160
208 aaa cag cag gca gct ctc cta cga ccc act gag gtc tac tgg gac ctg 825

```

## RAW SEQUENCE LISTING

DATE: 03/27/2002

PATENT APPLICATION: US/09/889,722A

TIME: 13:32:53

Input Set : A:\sequence listing.txt

Output Set: N:\CRF3\03272002\I889722A.raw

209	Lys	Gln	Gln	Ala	Ala	Leu	Leu	Arg	Pro	Thr	Glu	Val	Tyr	Trp	Asp	Leu	
210			165					170					175				
211	gac	atc	cag	acc	aat	gct	gtc	atc	aag	cac	cgg	ggg	cct	tca	gag	gtg	873
212	Asp	Ile	Gln	Thr	Asn	Ala	Val	Ile	Lys	His	Arg	Gly	Pro	Ser	Glu	Val	
213		180					185					190					
214	ctg	ccc	ccg	cat	ccc	gaa	gtg	gaa	ctg	ctc	cgc	tct	cag	ctc	atc	ctg	921
215	Leu	Pro	Pro	His	Pro	Glu	Val	Glu	Leu	Leu	Arg	Ser	Gln	Leu	Ile	Leu	
216	195					200				205						210	
217	aag	ctt	cgg	cag	cac	tat	cgg	gag	ctg	tgc	cag	cag	cga	gag	ggc	att	969
218	Lys	Leu	Arg	Gln	His	Tyr	Arg	Glu	Leu	Cys	Gln	Gln	Arg	Glu	Gly	Ile	
219					215					220					225		
220	gag	cct	cca	cgg	gag	tct	ttc	aac	cgc	tgg	atg	ctg	gag	cgc	aag	gtg	1017
221	Glu	Pro	Pro	Arg	Glu	Ser	Phe	Asn	Arg	Trp	Met	Leu	Glu	Arg	Lys	Val	
222			230					235					240				
223	gta	gac	aaa	gga	tct	gac	ccc	ctg	ttg	ccc	agc	aac	tgt	gaa	cca	gtc	1065
224	Val	Asp	Lys	Gly	Ser	Asp	Pro	Leu	Leu	Pro	Ser	Asn	Cys	Glu	Pro	Val	
225		245						250				255					
226	gtg	tca	cct	tcc	atg	ttt	cgt	gaa	atc	atg	aac	gac	att	cct	atc	agg	1113
227	Val	Ser	Pro	Ser	Met	Phe	Arg	Glu	Ile	Met	Asn	Asp	Ile	Pro	Ile	Arg	
228		260					265				270						
229	tta	tcc	cga	atc	aag	ttc	cgg	gag	gaa	gcc	aag	cgc	ctg	ctc	ttt	aaa	1161
230	Leu	Ser	Arg	Ile	Lys	Phe	Arg	Glu	Glu	Ala	Lys	Arg	Leu	Leu	Phe	Lys	
231	275				280					285						290	
232	tat	gcg	gag	gcc	gcc	agg	cgg	ctc	atc	gag	tcc	agg	agt	gca	tcc	cct	1209
233	Tyr	Ala	Glu	Ala	Ala	Arg	Arg	Leu	Ile	Glu	Ser	Arg	Ser	Ala	Ser	Pro	
234				295				300						305			
235	gac	agt	agg	aag	gtg	gtc	aaa	tgg	aat	gtg	gaa	gac	acc	ttt	agc	tgg	1257
236	Asp	Ser	Arg	Lys	Val	Val	Lys	Trp	Asn	Val	Glu	Asp	Thr	Phe	Ser	Trp	
237			310					315					320				
238	ctt	cgg	aag	gac	cac	tca	gcc	tcc	aag	gag	gac	tac	atg	gat	cgc	ctg	1305
239	Leu	Arg	Lys	Asp	His	Ser	Ala	Ser	Lys	Glu	Asp	Tyr	Met	Asp	Arg	Leu	
240			325					330					335				
241	gag	cat	ctg	cgg	agg	cag	tgt	ggc	ccc	cac	gtc	tcg	gcc	gca	gcc	aag	1353
242	Glu	His	Leu	Arg	Arg	Gln	Cys	Gly	Pro	His	Val	Ser	Ala	Ala	Ala	Lys	
243		340				345					350						
244	gac	tcc	gtg	gaa	ggc	atc	tgc	agt	aag	atc	tac	cac	atc	tcc	ctg	gag	1401
245	Asp	Ser	Val	Glu	Gly	Ile	Cys	Ser	Lys	Ile	Tyr	His	Ile	Ser	Leu	Glu	
246	355				360					365					370		
247	tac	gtc	aaa	cgg	atc	cga	gag	aag	cac	ctt	gcc	atc	ctc	aag	gaa	aac	1449
248	Tyr	Val	Lys	Arg	Ile	Arg	Glu	Lys	His	Leu	Ala	Ile	Leu	Lys	Glu	Asn	
249			375					380					385				
250	aac	atc	tca	gag	gag	gtg	gag	gcc	cct	gag	gtg	gag	ccc	cgc	cta	gtg	1497
251	Asn	Ile	Ser	Glu	Glu	Val	Glu	Ala	Pro	Glu	Val	Glu	Pro	Arg	Leu	Val	
252			390					395					400				
253	tac	tgc	tac	cca	gtc	cgg	ctg	gct	gtg	tct	gca	ccg	ccc	atg	ccc	agc	1545
254	Tyr	Cys	Tyr	Pro	Val	Arg	Leu	Ala	Val	Ser	Ala	Pro	Pro	Met	Pro	Ser	
255			405				410						415				
256	gtg	gag	atg	cac	atg	gag	aac	aac	gtg	gtc	tgc	atc	cgg	tat	aag	gga	1593
257	Val	Glu	Met	His	Met	Glu	Asn	Asn	Val	Val	Cys	Ile	Arg	Tyr	Lys	Gly	

VERIFICATION SUMMARY

DATE: 03/27/2002

PATENT APPLICATION: US/09/889,722A

TIME: 13:32:54

Input Set : A:\sequence listing.txt

Output Set: N:\CRF3\03272002\I889722A.raw

L:11 M:270 C: Current Application Number differs, Replaced Current Application Number

L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date